

CLAIMS

1. A method of executing queries on a data repository, the method comprising:
 - receiving a query, adapted for execution on a data set in the data repository;
 - defining a sample of the data set, wherein the sample is a subset of the data set;
 - executing the query on the sample;
 - generating an estimate of a result of the execution of the query on the sample; and
 - providing the estimate to a user interface.
2. A method of executing a query in accordance with claim 1, wherein the query includes criteria to provide the result of the execution of the query.
3. A method of executing a query in accordance with claim 1, wherein providing the estimate comprises displaying a representation of the estimate.
4. A method of executing a query in accordance with claim 1, the method further comprising:
 - defining an Nth sample of the data set, wherein the Nth sample is larger than an $(N - 1)$ th sample;
 - executing the query on the Nth sample;
 - generating an Nth estimate of the result based on the execution of the query on the Nth sample; and
 - providing the Nth estimate to a user interface.
5. A method of executing a query in accordance with claim 4, wherein the Nth sample of the data set is defined if the query is neither modified nor canceled after a preset time.
6. A method of executing a query in accordance with claim 4, wherein the Nth sample is defined to be larger than the $(N - 1)$ th sample by a factor Y.
7. A method of executing a query in accordance with claim 4, the method further comprising:

if the Nth sample is greater than or equal to a size Z,
executing the query on the data set to generate the result, and
providing the result to the user interface.

8. An information management system, the system comprising:
a data repository, wherein the data repository is configured to store a data set; and
a program for executing queries on the data repository, wherein the program is
operative to:

receive a query, adapted for execution on a data set in the data repository;
define a sample of the data set, wherein the sample is a subset of the data set;
execute the query on the sample;
generate an estimate of a result of the execution of the query on the sample;
and
provide the estimate to a user interface.

9. An information management system in accordance with claim 8, wherein the query
includes criteria to provide the result of the execution of the query.

10. An information management system in accordance with claim 8, wherein the
operation of providing the estimate of the result comprises displaying a representation of
the estimate.

11. An information management system in accordance with claim 8, wherein the program
is further operative to:

define an Nth sample of the data set, wherein the Nth sample is larger than an (N
– 1) th sample;
execute the query on the Nth sample;
generate an Nth estimate of the result based on the query of the Nth sample; and
provide the Nth estimate to a user interface.

12. An information management system in accordance with claim 11, wherein the Nth sample of the data set is defined if the query is neither modified nor canceled after a preset time.

13. An information management system in accordance with claim 11, wherein the Nth sample is defined to be larger than the $(N - 1)$ th sample by a factor Y.

14. An information management system in accordance with claim 11, wherein the program is further operative to:

if the Nth sample is greater than or equal to a size Z,
execute the query on the data set to generate the result, and
provide the result of the query execution to the user interface.